IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF WISCONSIN

WISCONSIN ALUMNI RESEARCH FOUNDATION,

Plaintiff,

v.

APPLE, INC.,

Defendant.

OPINION AND ORDER

14-cv-062-wmc

In this lawsuit, plaintiff Wisconsin Alumni Research Foundation ("WARF") alleges that defendant Apple, Inc. infringes U.S. Patent No. 5,781,752 (the "'752 patent"), which concerns a "table based data speculation circuit for parallel processing computer." Before the court are the parties' cross-motions for summary judgment and claim construction. (Dkt. ##116, 117.) For the reasons that follow, the court will adopt WARF's proposed construction of the term "prediction" and grant summary judgment to WARF on Apple's counterclaims and defenses based on anticipation under 35 U.S.C. § 102 with respect to U.S. Patent No. 5,619,662 ("Steely" or the "Steely patent"), as well as indefiniteness under 35 U.S.C. § 112 ¶ 2 with respect to claims 5 and 6. In turn, the court will deny Apple's motion for summary judgment based on those same defenses and counterclaims. As for Apple's motion for summary judgment on WARF's claim of willful infringement, the court will deny Apple's motion with respect to any defenses premised on (1) Apple's claim construction, (2) anticipation by Steely, and (3) indefiniteness of claims 5 and 6, but will reserve on Apple's motion in all other respects.

UNDISPUTED FACTS[1]

A. The Parties and Overview of This Lawsuit

Plaintiff Wisconsin Alumni Research Foundation ("WARF") is a Wisconsin corporation, with its principal place of business in Madison, Wisconsin. WARF is the owner of the '752 patent. Defendant Apple, Inc. is a California corporation, with its principal place of business in Cupertino, California.

On January 31, 2014, WARF filed suit against Apple alleging infringement of the '752 patent. Apple answered and asserted counterclaims for declaratory judgment of non-infringement and invalidity of the '752 patent. Material to the present motions, Apple contends that claims 1-3, 5, 6, and 9 of the '752 patent are invalid as anticipated by the "Steely patent. Apple also alleges that claims 5 and 6 of the '752 patent are invalid as indefinite.

B. Technology Overview

A modern computer device includes both hardware and software. Hardware typically includes memory, a microprocessor and peripherals, while software typically consists of sequences of instructions or "programs" that run on the hardware. At a general level, the microprocessor is responsible for fetching instructions and data, executing those instructions to modify the data, and then saving the results.[2] Typically,

[1] Except as otherwise noted, for purposes of summary judgment, the court finds the following facts to be material and undisputed.

[2] While the court uses the term "executing," the court acknowledges that the parties have agreed on a construction for the term "in fact executed" described below in subheading "E" of this Facts section.

individual instructions call for the performance of a relatively simple task, such as reading a value from or writing a value to a memory location, or adding, subtracting or comparing two numbers. There are generally three types of software instructions: (1) memory instructions; (2) computing instructions; and (3) control instructions.

Memory instructions are instructions that "when executed, cause data to be loaded into the processing unit from memory or stored from the processing unit to memory." ('752 Patent (dkt. #1-1) 1:38-40.) So-called "LOAD" instructions copy or read a value stored at a memory location specified by an address and return a value. LOAD instructions are also called "data consuming instructions," because they consume data by obtaining data from memory, though as Apple cautions, other types of software instructions also "consume" data. "STORE" instructions, on the other hand, copy or write a value to a memory location specified by an address. For that reason, STORE instructions are also called "data producing instructions," as they produce data by providing data to memory. (Apple similarly points out that other types of instructions "produce" data.) When a STORE instruction executes, it overwrites any value previously stored at that memory location. Both LOAD and STORE instructions are memory instructions.

Generally speaking, software instructions in a program have a predefined "program order," where the processor performs the instructions sequentially. Instructions, however, need not always be executed in the listed order. Instead, they may be executed "out of order." In out-of-order executions, instructions are typically executed when ready -- in other words, based on the availability of their input data, or "operands," rather than

a specified program order.[3] There are some obvious benefits to permitting instructions to execute out of order. For instance, because some instructions in a program take longer to execute than others, performing instructions in program order may slow processor performance since it requires waiting for earlier instructions to execute before performing later instructions in the program. Out-of-order execution, therefore, may result in increased efficiency since it allows the processor to use free time to execute other instructions that are ready to be processed. On the other hand, out-of-order execution may have a detrimental effect on performance if it leads to errors that require the processor to expend resources to correct.

A key requirement of efficient out-of-order execution, therefore, is that it must yield the same results as would the execution of instructions in program order. This requirement touches on the concept of "instruction dependency." A dependent instruction is one that must wait for the result of an earlier-in-order execution before it can safely execute.[4] For example, data dependency exists when an earlier-in-order STORE instruction writes data to the same address that is accessed by a later-in-order LOAD instruction. In that situation, the STORE and LOAD must execute in program order for the LOAD to read the correct data from the shared memory address that both instructions access.

[3] Apple clarifies that out-of-order executions also depend on the availability of necessary hardware.

[4] While "[a] processor may permit dependent instructions to execute out-of-order and then invoke a recovery process to return to a correct machine state," as Apple describes, Apple fails to dispute WARF's point that to execute "safely," the dependent instruction must wait to execute until after the instruction on which it depends has been executed. (Pl.'s Reply to Pl.'s PFOFs (dkt. #157) ¶ 21.)

In some situations, whether a given LOAD instruction depends on a STORE instruction from an earlier earlier-in-order program step cannot be known until after one or both of the instructions are executed. In other words, the processor lacks sufficient information to resolve whether or not a dependency actually exists. This uncertainty is known as "ambiguous dependency." Ambiguous dependencies may occur, for example, when the memory addresses that must be accessed by a given LOAD or STORE instruction are computed "on the fly" as the program executes. In those circumstances, the processor may have to perform additional computations with data that are not currently available in order to resolve whether one instruction is dependent on another.

To maximize processing speed, however, the processor may elect to execute a LOAD instruction before an earlier STORE instruction. The out-of-order execution of instructions without knowing if there is an actual dependency between them is known as "speculation" or "speculative execution," because the processor is speculating that there is no actual dependency. Speculation can be advantageous if it turns out to be correct (i.e., the LOAD instruction in fact was not dependent on the STORE instruction); then the out-of-order execution will yield the correct result and the performance will improve.[5] In contrast, if a LOAD instruction is speculatively executed ahead of a STORE instruction of earlier program order and it turns out that the speculation was incorrect (i.e., the

[5] The parties dispute whether "it is quite often the case that an ambiguous dependency is resolved as no dependency at all," as the '752 patent represents. ('752 patent (dkt. #1-1) 2:26-27.) As Apple contends, "[t]he degree to which ambiguous dependencies will turn out to be resolved as no dependency depends upon the workload." (Def.'s Resp. to Pl.'s PFOFs (dkt. #141) ¶ 30.) At the same time, Apple also proposes on summary judgment that "[m]any program instructions are 'independent' of each other and can safely execute-out-of-order with respect to each other." (Def.'s PFOFs (dkt. #119) ¶ 24.) Even if this could be construed as a dispute, it is not material to the issues before the court on summary judgment.

LOAD instruction was in fact dependent on the earlier STORE instruction), then the instructions will cause an error -- the prematurely executed LOAD instruction having obtained incorrect or stale data.[6]

In the patent-in-suit, this error is referred to as "mis-speculation," although the Steely patent -- as described below -- refers to it as a "collision." As discussed generally already, and as the patent-in-suit explains specifically, a mis-speculation can be detrimental to processor performance because it requires "the results of the prematurely executed dependent instructions [to] be discarded" or "squashed," and the instruction will need to be re-executed in program order. ('752 patent (dkt. #1-1) 2:46-49; Def.'s PFOFs (dkt. #119) ¶ 33.)[7]

C. The '752 Patent

i. Overview and Prosecution History

The '752 patent, entitled "Table Based Data Speculation Circuit for Parallel Processing Computer," was filed on December 26, 1996, and issued on July 14, 1998. The listed inventors are Drs. Andreas I. Moshovos, Scott E. Breach, Terani N. Vijaykumar, Gurindar S. Sohi. Plaintiff WARF is listed as the original assignee. WARF maintains that the named inventors conceived of the claimed invention no later than December 11, 1995.

[6] Apple maintains that there will be no error, at least technically, if the yet-to-execute STORE instruction would not change the value already written to the memory address, because the LOAD instruction still obtains the correct data. (Def.'s Resp. to Pl.'s PFOFs (dkt. #141) ¶ 31.)

[7] Apple contends that there may be times that the performance cost of mis-speculation does not outweigh the performance benefit of speculation. (Def.'s Resp. to Pl.'s PFOFs (dkt. #141) ¶ 34.)

During prosecution of the '752 patent, the named inventors provided no prior art to the Patent Office. On October 8, 1997, the patent examiner issued a Notice of References Cited, which listed four pieces of prior art. The patent examiner rejected pending claims 1-2, 6 and 8-11 as anticipated in light of U.S. Patent No. 5,555,432 ("Hinton"). On January 5, 1998, WARF filed a response cancelling pending claims 9 and 10, but arguing that claims 1, 2, 6, 8 and 11 were allowable over Hinton. On February 3, 1998, the examiner allowed those claims.

ii. Objectives and Specification

The '752 patent recognizes that "[t]he performance cost is a function of the frequency that speculation is required, the probability of mis-speculation and the time required to recover from a mis-speculation." ('752 patent at 3:14-18.) The '752 patent also observes that "most data dependent mis-speculations can be attributed to a few static STORE/LOAD instruction pairs," and that mis-speculations typically "exhibit 'temporal locality,'" such that "if one LOAD/STORE pair causes a data mis-speculation at a given point in time, it is highly likely that a later instance of the same pair will soon cause another mis-speculation." (*Id.* at 3:51-57.) The patent further observes that

> The present inventors believe that a relatively limited number of LOAD/STORE pairs will create mis-speculation and so the operation described above prevents the majority of the LOAD/STORE pairs from being slowed in execution. The list of critical LOAD/STORE pairs is prepared dynamically in a synchronization method for those LOAD/STORE pairs . . . .

(*Id.* at 14:15-22.) Based on these observations, the inventors concluded that load-based memory dependencies may be amenable to history-based prediction.[8] As such, the '752 patent associates predictions with particular LOAD instructions that have mis-speculated in the past.

The specification of the '752 patent describes a processor containing a "data speculation circuit" that detects dependence between LOAD and STORE instructions. The data speculation circuit also detects mis-speculations where a LOAD instruction that is dependent for its data on a STORE instruction appearing earlier in program order is in fact executed before the STORE instruction. According to the preferred embodiment, the data speculation circuit sends a "mis-speculation indication" to a "predictor circuit" if a mis-speculation is detected, which uses the indication to then produce a prediction. The greater the "prediction," the greater the likelihood that the speculative execution of its associated LOAD instruction will cause a mis-speculation; the lower a prediction at a given time, the lower the likelihood of mis-speculation. The processor uses each prediction to decide whether its associated LOAD instruction should be allowed to execute speculatively.

The patent discloses a "three-tiered approach" to dealing with ambiguous dependency. The first tier considers whether a LOAD instruction has a history of mis-speculation. "If there is no history of data mis-speculation, [the instruction] is executed without further inquiry." ('752 patent (dkt. #1-1) 3:64-66.) At this tier, the '752 patent

[8] Apple disputes that the named inventors were the first to develop history-based techniques for load-based memory dependencies. (Def.'s Resp. to Pl.'s PFOFs (dkt. #141) ¶ 42.)

describes a "prediction table," in which entries are created when the processor detects a mis-speculation by a LOAD instruction. "[I]f no entry is found in the prediction table," then "the reasonable assumption is that speculation can proceed." (*Id.* at 11:22-24.) The second tier becomes relevant with a LOAD instruction has mis-speculated in the past. In this tier, "a predictor based on the past history of mis-speculations for that LOAD instruction is employed to determine whether the instruction should be executed or delayed." (*Id.* at 4:1-4.) With respect to the second tier, the patent explains that "it is an object of the invention to provide a predictor circuit that may identify data dependencies on an on-going and dynamic basis." (*Id.* at 4:31-33.) Finally, in cases where the prediction indicates that the LOAD instruction should not be executed speculatively, the third tier may be employed to decide when the LOAD instructions should be allowed to execute. This part of the patent describes a "synchronization table," which "indicates whether there is in fact a pending LOAD instruction awaiting its dependent STORE instruction." (*Id.* at 11:45-47.)

iii. Claim Construction

a) "Prediction"

The heart of the parties' dispute turns on the meaning of the term "prediction" as used in claim 1 and all other independent claims. As context, WARF contends that the term "prediction" should be construed to mean "a variable that indicates the likelihood that the data speculative execution of a load instruction will result in a mis-speculation," where "a 'prediction' must be capable of receiving ongoing updates." (Pl.'s PFOFs (dkt. #122) ¶ 67.) In contrast, Apple contends that "prediction" need not be capable of

receiving updates, and it, therefore, proposes a construction of "a value that indicates that likelihood that the data speculative execution of a load instruction will result in a mis-speculation," but does not necessarily contemplate a revision to that value based on regular updates. (*Id.* at ¶ 68.)

b) Other Agreed-Upon Terms

The parties agree to the following constructions of claim terms:

- "data speculation circuit" (claims 1 and 9): "a circuit that detects data dependence between load and store instructions and that detects mis-speculation by load instructions"
- "mis-speculation" (claims 1, 6, and 9): "when a load instruction that is dependent for its data on a store instruction appearing earlier in the program order is in fact executed before the store instruction wrote its data to a memory address shared with the load instruction"
- "in fact executed" (claims 1 and 9): "when a load instruction has actually accessed a memory address that has not yet been updated by a store instruction appearing earlier in the program order"
- "predictor" (claim 1): "a circuit that receives a mis-speculation indication from the data speculation circuit to produce a prediction"

D. State of the Prior Art

i. Overview

By 1995, out-of-order processing was well-known in the field of computer architecture design. Also by 1995, techniques for detecting data dependence were well-known in the art. On this much, the parties are in agreement.

Apple further maintains that by 1995, data speculation was well-known in the art, as were techniques for detecting and recovering from mis-speculations. WARF disputes

this, asserting that the prior art techniques do not resemble the solutions proposed by the ’752 patent inventors. Apple also contends that by 1995, prediction techniques to improve the accuracy of speculation in an out-of-order processor were well-known in the art. WARF also disputes this, and in particular contends that the techniques disclosed in the prior art did not satisfy the “prediction” claimed in the ’752 patent -- the heart of the parties’ dispute addressed in the opinion below. Finally, Apple also maintains that by 1995, data speculation involving LOAD and STORE instructions was well-known in the art. WARF disputes that too, arguing the prior art techniques bore no resemblance to the solutions proposed in the ’752 patent.[9]

ii. The Steely Patent

The “Steely patent” is titled “Memory Reference Tagging” and names Simon C. Steely, Jr., David J. Sager and David B. Fite, Jr. as inventors. The application was filed on August 12, 1995, and claims priority to an earlier application filed on November 12, 1992. The Steely patent issued as U.S. Patent No. 5,619,662 on April 8, 1997, and was assigned to DEC. As such, it is prior art to the ’752 patent. Apple contends that the Steely patent anticipates claims 1-3, 5, 6, and 9 of the ’752 patent.

[9] Apple proposes finding of facts about other prior art references, including: a technique developed by Digital Equipment Corporation (“DEC”); U.S. Patent No. 5,666,506 (“Hesson”); and a commercial processor known as the Alpha 21264 or “EV6.” As best as the court can discern, however, these prior art references are only material to Apple’s motion for summary judgment on the objective prong of WARF’s willful infringement claim. As discussed below, the court reserves on that based on any arguments not developed fully at summary judgment, waiting instead to hear the evidence of infringement and invalidity to be introduced during the first phase of the trial.

Pertinent to this anticipation defense, all claims of the '752 patent require a "prediction" associated with a LOAD instruction or with a LOAD/STORE pair. Apple maintains that the Steely patent describes a processor that executes instructions out of order and uses a prediction to determine whether to allow speculation for LOAD and STORE instructions. WARF asserts that Steely fails to disclose any "prediction" capable of receiving ongoing updates -- or even a "prediction" under Apple's proposed construction of that claim term.

E. Person of Ordinary Skill in the Art

The parties dispute what characteristics a person of ordinary skill in the art would possess, though this dispute is not material to the parties' respective motions for summary judgment, or at least the reasons for this court's disposition of those motions. Apple maintains that for purposes of the '752 patent, a person of ordinary skill in the art would have a Ph.D. in electrical engineering, computer engineering or computer science with a focus on computer architecture or microprocessor design; or an M.S. or B.S. degree in electrical engineering, computer engineering or computer science with significant work experience relating to computer architecture or microprocessor design.

WARF maintains that a person of ordinary skill in the art would have at least a bachelor's degree in electrical engineering or computer science, and at least three to five years of experience in computer design and computer architecture. Alternately, WARF asserts a person of ordinary skill in the art would have a master's degree in electrical engineering or computer science, and at least two to three years of experience in

computer design and computer architecture. The experience could be derived from either industry or academia.

F. IPR Decision

Finally, Apple filed a petition with the United States Patent and Trademark Office's Patent Trial and Appeal Board ("PTAB") seeking *inter partes* review ("IPR") of all claims of the '752 patent. In the petition, Apple argued that claims 1-9 are invalid as obvious in view of Hesson and Steely, relying on a declaration of its expert Dr. Colwell. On April 15, 2015, after briefing by Apple and WARF, the PTAB denied Apple's petition "as to all challenged claims," finding that Apple "has not shown . . . that there is a reasonable likelihood that it will prevail" on its obviousness theory for any claim of the '752 patent. (4/17/2015 Declaration of Christopher Abernathy ("4/17/15 Abernathy Decl."), Ex. A (dkt. #151-1) p.3.)

In particular, the PTAB construed "prediction" as "a variable that indicates the likelihood that the data speculative execution of a load instruction will result in a mis-speculation." (*Id.* at p.10.) In so finding, the PTAB reasoned:

> We agree that in the '752 patent, the mis-speculation prediction at any point in time is a function of the mis-speculation history of load-store instruction pairs. Thus, the prediction is a variable. The fact that the prediction has a particular value at each point in time is merely an indication of its functional relationship and does not change the nature of the prediction from a variable to a constant value.

(*Id.*)[10]

[10] In *Wisconsin Alumni Research Foundation v. Intel Corp.*, No. 08-cv-78-bbc (W.D. Wis. filed Feb. 5, 2008), Judge Crabb similarly construed "prediction" in the same patent to mean "a variable that

OPINION

I. Claim Construction

"It is a 'bedrock principle' of patent law that 'the claims of a patent define the invention to which the patentee is entitled the right to exclude.'" *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312 (Fed. Cir. 2005) (en banc) (quoting *Innova/Pure Water, Inc. v. Safari Water Filtration Sys., Inc.*, 381 F.3d 1111, 1115 (Fed. Cir. 2004)). The court exclusively determines claim construction as a matter of law. *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 372 (1996). The words of the claims are always the "appropriate starting point" for proper construction, *Comark Commc'ns, Inc. v. Harris Corp.*, 156 F.3d 1182, 1186 (Fed. Cir. 1998), with the court asking "how a person of ordinary skill in the art understands a claim term" as an "objective baseline from which to begin claim interpretation," *Phillips*, 415 F.3d at 1313.

"Importantly, the person of ordinary skill in the art is deemed to read the claim term not only in the context of the particular claim in which the disputed term appears, but in the context of the entire patent, including the specification." *Id.* In fact, "[t]he best source for understanding a technical term is the specification from which it arose, informed, as needed, by the prosecution history." *Multiform Desiccants, Inc. v. Medzam, Ltd.*, 133 F.3d 1473, 1478 (Fed. Cir. 1998). As the Federal Circuit has recognized,

indicates the likelihood that the data speculative execution of a load instruction will result in mis-speculation," and later clarified that "a 'prediction' must be capable of receiving ongoing updates." *Wis. Alumni Research Found. v. Intel Corp.*, No. 08-cv-78-bbc, 2008 WL 4279975, at *7 (W.D. Wis. Sept. 18, 2008); *Wis. Alumni Research Found. v. Intel Corp.*, 656 F. Supp. 2d 898, 922 (W.D. Wis. 2009). While the court agrees with Apple that this decision has no binding effect on this court (Def.'s Opening Br. (dkt. #118) 23), any more than the PTAB's decision does, the court will obviously consider Judge Crabb's and the PTAB's reasoning and analysis in the opinion below.

however, "there is sometimes a fine line between reading a claim in light of the specification, and reading a limitation into the claim from the specification." *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 904 (Fed. Cir. 2004) (quoting *Comark Commc'ns*, 156 F.3d at 1186-87). "[A]n inherent tension exists as to whether a statement is a clear lexicographic definition or a description of a preferred embodiment. The problem is to interpret claims 'in view of the specification' without unnecessarily importing limitations from the specification into the claims." *E-Pass Techs., Inc. v. 3Com Corp.*, 343 F.3d 1364, 1369 (Fed. Cir. 2003).

In addition to intrinsic evidence like the specification and prosecution history, the Federal Circuit has "authorized district courts to rely on extrinsic evidence, which 'consists of all evidence external to the patent and prosecution history, including expert and inventor testimony, dictionaries, and learned treatises.'" *Phillips*, 415 F.3d at 1317 (quoting *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 980 (Fed. Cir. 1995)). "However, while extrinsic evidence 'can shed useful light on the relevant art,' [the Federal Circuit has] explained that it is 'less significant than the intrinsic record in determining 'the legally operative meaning of claim language.'" *Id.* (quoting *C.R. Bard, Inc. v. U.S. Surgical Corp.*, 388 F.3d 858, 862 (Fed. Cir. 2004)). Accordingly, the court can consider extrinsic evidence in construing patent claims, but it must do so in the context of the intrinsic evidence and while keeping in mind the flaws inherent in each type of extrinsic evidence. *Id.* at 1318.

As previously mentioned, the parties dispute the proper construction of only one claim term, "prediction," which appears in claims 1, 2, 3, 5 and 9 of the '752 patent.[11] The parties propose the following constructions for that term:

"Prediction"	
Plaintiff WARF's Proposed Construction	**Defendant Apple's Proposed Construction**
"A variable that indicates the likelihood that the data speculative execution of a load instruction will result in a mis-speculation" A "prediction" must be capable of receiving ongoing updates.	"A value that indicates the likelihood that the data speculative execution of a load instruction will result in a mis-speculation" A "prediction" need not be capable of receiving ongoing updates.

The obvious and sole substantive difference between the parties' competing constructions is whether the prediction *must* be capable of change (while implicit in the "value"-"variable" dispute, the second sentence of each definition makes that disagreement explicit). According to WARF, a "prediction" must be able to receive updates -- in other words, it must be dynamic. Apple, on the other hand, argues that a "prediction" *may* be dynamic, but it may *also* be static -- that is, incapable of receiving ongoing updates and changing to reflect those updates.

The language of claim 1 reads as follows:

> In a processor capable of executing program instructions in an execution order differing from their program order, the processor further having a data speculation circuit for detecting data dependence between instructions and detecting a mis-speculation where a data consuming instruction dependent for its data on a data producing

[11] The parties also mention a possible dispute about the meaning of "table," but neither party sought construction at summary judgment.

> instruction of earlier program order, is in fact executed before the data producing instruction, a data speculation decision circuit comprising:
>
> a) a predictor receiving a mis-speculation indication from the data speculation circuit to produce a *prediction* associated with the particular data consuming instruction and based on the mis-speculation indication; and
>
> b) a prediction threshold detector preventing data speculation for instructions having a *prediction* within a pre-determined range.

('752 patent, 14:36-52 (emphasis added).)

The court can dispense with one of Apple's arguments at the outset. Apple points out that none of the claims at issue contain an *express* limitation requiring the prediction to be updated on an ongoing basis, suggesting that this means predictions need not be capable of update. (Def.'s Br. Support Summ. J. (dkt. #118) 19-20.) As appealing as that simple construction might be, since it would alleviate the need for further analysis, the lack of an express limitation actually requires further inquiry: if the claims contained an additional limitation requiring a *dynamic* prediction, construing the term "prediction" to be intrinsically dynamic would render that limitation superfluous. *See Digital-Vending Servs. Int'l, LLC v. Univ. of Phoenix, Inc.*, 672 F.3d 1270, 1275 (Fed. Cir. 2012) (discussing the "well-established rule that claims are interpreted with an eye toward giving effect to all terms in the claim") (internal quotation marks and citation omitted); *cf. LSI Indus., Inc. v. ImagePoint, Inc.*, 279 F. App'x 964, 972 (Fed. Cir. 2008) ("Some claims specifically recite 'an illuminated display device,' while others recite only 'a display device.' . . . Thus, the language of the claims counsels against imposing an illumination limitation on the display device term because it would make the limitation superfluous where it explicitly

appears."); *Phillips*, 415 F.3d at 1314 ("To take a simple example, the claim in this case refers to 'steel baffles,' which strongly implies that the term 'baffles' does not inherently mean objects made of steel."). Thus, although the claims include no limitations explicitly requiring predictions to be dynamic, the word "prediction" itself still might (or might not) include that requirement depending on the claim language, specification, prosecution history and extrinsic evidence.

WARF relies heavily on the fact that the claimed "data speculation decision circuit" prevents data speculation for instructions having a prediction "within a pre-determined range." According to WARF, the claimed function of determining whether a prediction falls within a given range makes sense only if, "at any given time *after* the 'prediction' is produced, it *might* be 'within a predetermined range' or it might *not* be." (Pl.'s Br. Opp'n Summ. J. (dkt. #148) 8 (emphasis in original).) Apple argues in response that this interpretation narrows the claims in a way not supported by the text or the state of the prior art. In particular, Apple points out that prior art in the field, including an article entitled "A Study of Branch Prediction Strategies" by James E. Smith (Decl. of Bryan S. Conley, Ex. 12 (dkt. #124-12) [hereinafter "Smith" or the "Smith article"]), used the word "prediction" in the context of speculation strategies tracking single past events, rather than a dynamic history of such events.[12] According to Apple,

[12] The Smith article deals with control speculation, rather than data speculation. Control speculation involves "branch prediction." In the words of the '752 patent, it "might involve executing an instruction that follows a branch instruction without knowing the outcome of the branch (and thus whether the following instruction should have been executed or was branched around)." ('752 patent, 2:32-36.)

the claims certainly *permit* a dynamic prediction but are also broad enough to encompass a prediction incapable of receiving updates.

In *Wisconsin Alumni Research Foundation v. Intel Corp.*, 656 F. Supp. 2d 898 (W.D. Wis. 2009), this court relied in part on the same language WARF cites, finding that:

> Claim 1 describes "producing" a "prediction" from a "mis-speculation indication" generated in a data speculation circuit and determining whether that "prediction" is "within a predetermined range" to decide whether to prevent data speculation. Thus, the claim language itself establishes that a "prediction" is something other than a stored "indication" and is capable of having a "range" of values . . . On its face, this language suggests that a prediction must be capable of changing over time.

Id. at 922.

Revisiting this same claim language here, the court again finds the contemplated use of a predetermined range of values to assess whether instructions should be permitted to speculate favors WARF's narrower interpretation. By way of example, imagine a pair of instructions that mis-speculates for the first time. The parties agree that a single mis-speculation is enough to produce a prediction; thus, in this instance, the predictor of the invention would receive that mis-speculation indication from the data speculation circuit and use it to produce a prediction of "1," representing the single mis-speculation. Under Apple's construction, the development of the prediction can end here, because it need not be capable of further updates. Thus, the prediction would be set permanently at its initial value of "1."

Under this approach, the next time the instructions execute, there is no need for the prediction threshold detector to assess whether the prediction falls within a "pre-

determined range." In a static situation, there are only two possibilities: either there *is* a prediction with a value of "1," because the instructions have mis-speculated a single time; or there is *no* prediction, because the instructions have not yet mis-speculated and, therefore, the prediction has not yet been created. Thus, under Apple's construction, the question for the prediction threshold detector is a binary determination of whether a prediction *exists at all*, rather than whether a prediction "falls within a given range." Indeed, there would be no need for "a data speculation *decision* circuit" in claim 1 at all, feeding ongoing mis-speculation outcomes, since the "data speculation circuit" itself would provide the single piece of information required for a static prediction.

Effectively, the prediction threshold detector would prevent data speculation.[13] Said another way, Apple's construction would read out the words "within a pre-determined range" from subsection (b) of claim 1, or at least render them superfluous in the context of "predictions" incapable of receiving updates; in those cases, the prediction threshold detector would prevent data speculation "for instructions having a prediction," full stop.

Of course, as Apple argues, a "range" can consist of a single value, which could technically allow for a "predetermined range" including *only* the value 1. Superficially, this provides some support for Apple's construction, but it still does not explain *why* it would ever be necessary to compare an existing prediction to a range of 1 for so-called

[13] Theoretically, it is possible that the prediction of "1" would *not* fall within the predetermined range and the instruction would be allowed to execute regardless of the previous mis-speculation. But if that were so, the invention would appear to serve no purpose, since the prediction would not prevent speculation and could not change, much less improve, the processor's performance over time. Likely for this reason, Apple does not advance this argument, so the court does not consider it further.

static "predictions." The choice is still binary -- either there is no prediction or the prediction is set to 1 -- and so the notion of "comparison" remains a poor fit for the kind of theoretical static "predictions" Apple posits, regardless of whether the predetermined range is set to encompass multiple values or a single value.

The remainder of the '752 patent further supports WARF's construction. The brief summary of the invention describes a three-tiered approach for determining when an instruction should execute. The first tier encompasses instructions with no history of mis-speculation; they may execute "without further inquiry." ('752 patent, 3:66.) The second tier implicates instructions that have previously mis-speculated. At that point, according to the description, the invention employs a predictor "to determine whether the instruction should be executed or delayed." (*Id.* at 4:1-4.) If the prediction were static, however, the mere fact of its existence would be enough to prevent execution. In contrast, the predictor as described in the '752 patent instead uses "the past history of mis-speculations" to determine whether the instruction may execute, allowing those that are "*typically* not dependent" to execute immediately. (*Id.* at 4:1-5 (emphasis added).) This language, too, suggests a prediction capable of update; it makes little sense to speak of instructions that are "typically not dependent" when a single instance of mis-speculation could, under Apple's construction, foreclose future speculative execution without the possibility of updates to reflect what *typically* occurs. If the predictor ultimately delays the instruction, the third tier then employs a synchronization table to determine when the instruction should execute, delaying it "until after the execution of the particular data producing instruction" on which it depends. (*Id.* at 4:5-7, 27-28.)

Furthermore, this three-tiered approach appears in the brief summary of the invention, rather than as a description of a single embodiment, making it more persuasive as a source of support for WARF's narrower construction. *See C.R. Bard, Inc. v. U.S. Surgical Corp.*, 388 F.3d 858, 864 (Fed. Cir. 2004) ("Statements that describe the invention as a whole, rather than statements that describe only preferred embodiments, are more likely to support a limiting definition of a claim term. . . . Statements that describe the invention as a whole are more likely to be found in certain sections of the specification, such as the Summary of the Invention.").

While less persuasive given its location in the patent, the detailed description of the invention provides further context suggesting that a "prediction" must be dynamic. As this court recognized in describing the preferred embodiment of the invention in *Intel*, "the specification explains in unequivocal terms that '[t]he prediction provided by the predictor circuit 33 . . . is updated based on historical mis-speculations detected by the data speculation circuit 30. For this reason, the data speculation circuit 30 must communicate with the predictor circuit 33 on an *ongoing* basis.'" *Intel*, 656 F. Supp. 2d at 922 (quoting '752 patent, 8:7-11) (emphasis added). WARF also points to other examples supporting its position in the description of the preferred embodiment, including the description of the way the prediction normally "is *incremented and decremented*" such that "the higher the prediction **109**, the more likelihood of mis-speculation[.]" ('752 patent, 11:29-35 (emphasis added).) Indeed, throughout the description of the preferred embodiment, the specification consistently refers to the prediction as dynamic. (*See, e.g.*, *id.* at 12:14-17 ("[T]he prediction that there was a need

to synchronize was wrong and so at process block **120** the prediction **109** is *decremented* toward the do not synchronize state.") (emphasis added); 12:52-54 ("[T]he prediction **109** is *updated* toward the synchronize condition indicating that the prediction that there was a need to synchronize was correct[.]") (emphasis added); 12:67-13:3 ("If [a mis-speculation occurs and the pair is already in the prediction table] then at process block **302**, the prediction **109** is *updated* toward synchronize so that this mis-speculation may be avoided in the future.") (emphasis added).)

Acknowledging, as it must, that the preferred embodiment describes a dynamic prediction that receives ongoing updates (Def.'s Br. Support Summ. J. (dkt. #118) 21; Def.'s Br. Opp'n Summ. J. (dkt. #140) 16), Apple relies on the general principle that "it is improper to read limitations from a preferred embodiment described in the specification -- even if it is the only embodiment -- into the claims absent a clear indication that the patentee intended the claims to be so limited." *Liebel-Flarsheim Co.*, 358 F.3d at 913. But this is not a case in which the "claim language is sufficiently broad that it can be read to encompass features not described in the written description, either by general characterization or by example in any of the illustrative embodiments." *Id.* at 905. Rather, as described above, the claims *themselves* suggest that the contemplated "prediction" must be capable of change; the preferred embodiment merely provides further support for that conclusion.

Use of the preferred embodiment as context, rather than as a source of limitations that do not otherwise appear in the claims, is permissible. *Compare, e.g.*, *Teleflex Inc. v. Ficosa N. Am. Corp.*, 299 F.3d 1313, 1327-28 (Fed. Cir. 2002) (district court erred in

holding that "clip" was limited to a "single pair of legs," even where that was the only embodiment described, where claim language did not support that limitation, specification and prosecution history included no statements of restriction and the ordinary meaning of "clip" was not so restricted), *with Toro Co. v. White Consolidated Indus., Inc.*, 199 F.3d 1295, 1301-02 (Fed. Cir. 1999) (holding that construction of "including" required attachment between structures where that was the only embodiment disclosed *and* where nothing in the remainder of the specification supported an unattached embodiment; "[T]he specification describes the advantages of the unitary structure as important to the invention. . . . No other, broader concept was described as embodying the applicant's invention, or shown in any of the drawings, or presented for examination.").

Still, Apple argues that the '752 patent *does* expressly contemplate alternative embodiments of the invention, pointing out that the detailed description of the invention states:

> It will be understood that the prediction **109** may be obtained by methods other than simply incrementing it in value for each speculation as described herein. For example, various weighting schemes can be provided to cause the predictor circuit **33**, for example, to be less sensitive to the earliest mis-speculations. More complex pattern matching techniques may also be used, for example, to catch situations where mis-speculations occur in groups or regular patterns.

('752 patent, 14:6-14.) Apple contends that a person of ordinary skill in the art would read this discussion to allow for alternative embodiments in which a prediction is *not* updated on an ongoing basis and urges the court not to "improperly exclude a disclosed embodiment" by adopting WARF's construction. *See Broadcom Corp. v. Emulex Corp.*, 732

F.3d 1325, 1333 (Fed. Cir. 2013). The problem with this argument is that the patent simply does *not* disclose the embodiment Apple advocates.

As a beginning point, *neither* of the alternative embodiments disclosed in the specification contemplate a static "prediction." To the contrary, both proposed alternatives implicitly contemplate arrangements involving dynamic predictions. Both schemes that assign a different weight to later mis-speculations and techniques that identify mis-speculations occurring in groups or regular patterns assume a *developing history* of mis-speculations that the predictor circuit can use to obtain its prediction. There is no need to weight different instances of mis-speculation if the prediction is static and will never be updated to reflect those weights.[14] There is likewise no need to develop complex matching techniques to identify patterns in mis-speculation if the prediction can never take that information into account in determining how likely an instruction is to mis-speculate. Adopting WARF's construction, therefore, does not exclude a "disclosed embodiment" from the scope of the claims.

Nor is there support for Apple's proposed construction in the intrinsic evidence, Apple's arguments to the contrary notwithstanding. For instance, Apple argues that the specification makes clear that a single mis-speculation is enough to produce a prediction. (Def.'s Br. Support Summ. J. (dkt. #118) 18.) This is true enough, but it has no bearing

[14] Apple uses the mention of "weighting schemes" to propose its own take on an alternative embodiment as well -- a "weighting scheme that always prevents speculation by a load instruction for which mis-speculation recovery would be especially costly." (Def.'s Br. Support Summ. J. (dkt. #118) 23.) Such a weighting scheme would *not*, however, be a means of producing a "prediction." Both parties agree that a prediction indicates the *likelihood* that a pair of instructions will mis-speculate; Apple's embodiment has nothing to do with the likelihood of mis-speculation, but rather assesses whether the costs associated with a single mis-speculation are prohibitive, regardless of how likely or unlikely that mis-speculation is.

on whether a prediction must be capable of update *after* its creation -- nor does the fact that the claim language "does not specify any minimum number of times that the instruction must mis-speculate before the 'prediction' is above the threshold required to prevent speculation." (*Id.*) Apple also argues that a "prediction" must be construed as a "value" (i.e., static) rather than a "variable" (i.e., dynamic) because the specification "explicitly describes the prediction as a 'value.'" (*Id.* at 15-16.) But the examples it cites speak of the prediction being set to a "default value" or being "incremented in value." All this confirms is that the prediction is some number that *has* a value; it does not suggest the value of that prediction cannot change. To the contrary, the portions of the specification Apple cites refer to "incrementing" the value of the prediction, suggesting that it can and does change. Thus, the court again adopts Judge Crabb's conclusion in *Intel* that "[n]either the claim language nor the specification supports defendant's proposed construction that a 'prediction' may include values that are fixed once to indicate a single incident of mis-speculation." 656 F. Supp. 2d at 922.

Finally, Apple contends that its construction finds support in *extrinsic* evidence, citing to the Smith article discussed above, as well as the reports of its two experts, Dr. August and Dr. Colwell. Both reports, however, primarily rehash Apple's legal arguments by purporting to analyze the language in the specification and claims. (*See* August Report (dkt. #103) ¶¶ 136-47; Colwell Report (dkt. #104) ¶¶ 131-41.) The court rejects Apple's positions on those issues, and so, too, expert reports that echo those same arguments.

Apple is, therefore, left with the Smith article and another paper, "Memory Dependence Prediction [U]sing Store Sets," by George Z. Chrysos and Joel S. Emer (the "Chrysos paper") (Conley Decl., Ex. 1 (dkt. #143-1)), both of which Apple contends describe techniques that produce static predictions. Even if Apple's characterization were accurate, these two extrinsic references are wholly underwhelming compared to the language of the patent itself and contrary intrinsic evidence. Moreover, extrinsic evidence "can be used only to help the court come to the proper understanding of the claims"; it cannot be used to vary or contradict the claim language or specification. *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1584 (Fed. Cir. 1996).

For all these reasons, the court finds WARF's proposed construction of the term "prediction" compelling and will construe that term as requiring a prediction that is capable of receiving updates.

II. Invalidity

A. Anticipation by Steely

On summary judgment, both parties devote most of their invalidity briefing to the question of whether the '752 patent is invalid as anticipated by the Steely patent, U.S. Patent No. 5,619,662. Evaluating a claim of anticipation involves a two-step inquiry. The first step requires proper construction of the meaning and scope of the claims. *Power Mosfet Techs., L.L.C. v. Siemens AG*, 378 F.3d 1396, 1406 (Fed. Cir. 2004). "The second step in the analysis requires a comparison of the properly construed term to the prior art[.]" *Id.* To demonstrate anticipation, "the proponent must show 'that the four corners of a single, prior art document describe every element of the claimed invention.'"

Net MoneyIN, Inc. v. VeriSign, Inc., 545 F.3d 1359, 1369 (Fed. Cir. 2008) (quoting *Xerox Corp. v. 3Com Corp.*, 458 F.3d 1310, 1322 (Fed. Cir. 2006)). Although anticipation is ultimately a question of fact, "it may be decided on summary judgment if the record reveals no genuine dispute of material fact." *Leggett & Platt, Inc. v. VUTEk, Inc.*, 537 F.3d 1349, 1352 (Fed. Cir. 2008) (quoting *Golden Bridge Tech., Inc. v. Nokia, Inc.*, 527 F.3d 1318, 1321 (Fed. Cir. 2008)).

Both parties have moved for summary judgment of anticipation in light of Steely, and each relies on its own, preferred construction of the disputed term "prediction." (Def.'s Br. Support Summ. J. (dkt. #118) 25-37; Pl.'s Br. Support Summ. J. (dkt. #120) 38-64.) Having just rejected Apple's construction of that term, Apple's motion for summary judgment will be denied. Even if the court adopts WARF's construction, however, Apple maintains numerous disputed issues of fact preclude entry of summary judgment against it on grounds of anticipation. (Def.'s Br. Opp'n Summ. J. (dkt. #140) 27.) It is to that question the court now turns.

i. Background of the Steely Patent

The Steely patent is entitled "Memory Reference Tagging" and describes a processor that "includes a memory reference tagging store associated with the instruction scheduler so that the scheduler can reorder memory reference instructions without knowing the actual memory location addressed by the memory reference instruction." (U.S. Patent No. 5,619,662 (dkt. #131-4) Abstract.) Most relevant to the issue of anticipation, Steely discloses four different techniques in which a "write buffer" assigns "memory reference tags" involving a mis-speculation to load and store instructions. Each

of those techniques appears in the section of the patent entitled "Memory Reference Tagging." (*See id.* at 47:35-49:8.)

In the first technique, a mis-speculation generates a memory reference tag from a portion of the address in memory that resulted in the LOAD-STORE collision. (*Id.* at 48:2-4.) Once that portion is placed in the memory tag store, every time an instruction is retrieved from memory to be executed, the memory reference tag circuit "will provide the tag bits to be used by the instruction scheduler." (*Id.* at 48:30-33.) If the instructions appear with identical tag bits (indicating a previous mis-speculation), the instruction scheduler will not reorder them. (*Id.* at 48:33-36.)

In the second technique, the pair of instructions after a mis-speculation is tagged not with a portion of the memory address, but instead with "a problem number which could be a number provided from a counter." (*Id.* at 48:55-57.) As a result, "[t]wo memory reference instructions with the same address and number will not reorder." (*Id.* at 48:57-59.) "However, if the two memory reference instructions have a different number, the instructions will reorder." (*Id.* at 48:59-61.) The counter does not appear to increment with respect to that for the same pair of instructions once it has assigned the "problem number"; rather, it increments when a mis-speculation occurs with respect to a *different* pair of instructions. For instance, if a pair of instructions mis-speculates and is assigned the problem number 0, the next pair to mis-speculate might be assigned the problem number 1.

The third technique is to assign an instruction a "bit to indicate that an instruction should not be reordered." (*Id.* at 48:62-63.) Thus, using this technique, "for

a store that previously caused a problem in the write buffer, the instruction is tagged with a bit indicating that the ISCHED **38** [instruction scheduler] cannot reorder memory reference instructions around the Instruction tagged with the bit." (*Id.* at 48:63-67.)

The final technique is to "turn off reordering" entirely in response to a mis-application under certain circumstances. (*Id.* at 49:4-5.) For example, the patent suggests turning off reordering when entering a subroutine, based on the general observation that "during a subroutine call, there are some initial stores and some exiting loads" and "[i]t would not be desirable to reorder the exiting loads before the initial stores." (*Id.* at 49:4-9.)

ii. Analysis

Construing "prediction" as a dynamic (updating) "variable that indicates the likelihood that the data speculative execution of a load instruction will result in a mis-speculation," the remaining question for deciding the anticipation issue before the court is whether Steely discloses a prediction that can change over time. In *Intel*, this court found that it did not, holding that the requirement of a dynamic prediction was "fatal to defendant's contention that the four techniques disclosed in the '662 patent anticipate the '752 patent." 656 F. Supp. 2d at 922. WARF urges the court to adopt the same result here, arguing that Steely's four techniques do not disclose predictions that update on an ongoing basis. According to WARF, those techniques simply involve tagging instructions to reflect a single mis-speculation event, without providing a mechanism to update those tags. *See also Intel*, 656 F. Supp. 2d at 922 ("For each [technique], the tag is

designed to indicate only that a mis-speculation has occurred, not keep track of mis-speculations on an ongoing basis.").

Unsurprisingly, Apple objects to this characterization. Apple instead contends that the outcome of Steely's tag comparison "*can* change over time for the same pair of load and store instructions." (Def.'s Br. Opp'n Summ. J. (dkt. #140) 29.) According to Apple's expert, Dr. Colwell, that can occur if, after the write buffer assigns the tags, additional mis-speculations involving one of the pair of tagged instructions occur. As an example, Dr. Colwell presumes a situation in which a load instruction, "Inst 1011," and store instruction, "Inst 1007," have been tagged with the same memory address of "10010," such that Steely would prevent speculation. (Colwell Report (dkt. #104) ¶ 303.) Colwell then posits another situation in which a *different* load instruction, Inst 1012, *also* mis-speculates with store instruction Inst 1007:

> Another load instruction, for instance Inst 1012, may later be reordered ahead of store instruction Inst 1007, both Inst 1012 and Inst 1007 accessing the same memory address, this address ending in a different set of 5 bits, for instance "00110." Inst 1007 would then be associated with the tag "00110," which would no longer be identical to the tag "10010" associated with the load instruction Inst 1011. Because the tags for Inst 1007 and Inst 1011 are no longer identical, Steely predicts they are not dependent and may reorder them. Thus, the "prediction" disclosed by Steely is a "variable" that is "capable of receiving ongoing updates," as required by WARF's proposed construction of the term "prediction."

(*Id.*)

Whether Steely actually discloses this means of "updating" its tags within the four corners of the patent is certainly open to debate. Apple asks the court to infer as much,

based on the fact that: (1) the memory reference tag store is large enough to store just one tag per instruction; and (2) the patent describes how tags for each mis-speculation "will be stored" regardless of other tags that may already exist for those instructions. According to Apple, these two facts demonstrate that Steely overwrites previously stored tags, thereby "updating" the result of any comparison that Steely performs between the two.

The flaws in this argument are multiple. Essentially, Apple and its experts assume what amounts to a defect in Steely, which prevents it from assigning more than one tag preventing mis-speculation to a single instruction (in Dr. Colwell's example, Inst. 1007), even though the example posits tags with different memory addresses (Inst. 10010 and Inst. 00110), depending on the store instruction with which the load instruction 1007 is paired (here, Inst. 1011 and Inst. 1012). Not only is this assumption contradicted by the language in Steely, *see* U.S. Patent no. 5,619,662, at 47:37-43 ("The memory reference tag store . . . provides *at least* one bit associated with said instruction . . .") (emphasis added), but it would undermine the whole purpose of Steely, which is to prevent future mis-speculations, since it would result in a never-ending loop for load instructions causing multiple mis-speculations each time the 10010 and 00110 tags overwrite one another.[15]

[15] The inventor of Steely does appear to have admitted in his deposition that in his view, this is how his invention would function, though he was asked the question out of context and without being asked about the obvious defect this would appear to create in his patent. (*See* Def.'s Br. Opp'n Summ. J. (dkt. #140) 31.) As WARF points out, this is why after-the-fact testimony of the inventor is of limited relevance when unsupported by the patent itself. *See Howmedica Osteonics Corp. v. Wright Med. Tech., Inc.*, 540 F.3d 1337, 1346 (Fed. Cir. 2008) ("The testimony of an inventor 'cannot be relied on to change the meaning of the claims.'") (quoting *Markman*, 52

More importantly for summary judgment purposes, even assuming one might infer that such overwriting occurs and constitutes "updating," the above-described "tag replacement system" would hardly constitute a "prediction" as this court has construed the term. Properly construed, a "prediction" must communicate the likelihood of mis-speculation *and* must be capable of update. Using the example offered by Dr. Colwell for the sake of simplicity, Inst 1011 and Inst 1007 proved to be dependent and were, accordingly, tagged with the same memory address. Thereafter, *another* load instruction, Inst 1012, also proves to be dependent on store instruction Inst 1007. Accordingly, Steely overwrites the first tag on Inst 1007, tagging it to match Inst 1012, but that is not so much an "update" of the comparison between Inst 1007 and Inst 1011 as it is the wholesale *elimination* of that comparison. By Apple's and Dr. Colwell's own description, *no* record of the previous mis-speculation remains; the next time the tags are compared under Steely, they fail to reflect that any mis-speculation has occurred in the past and, therefore, fail to communicate the likelihood that the data speculative execution of the load instruction 1011 and store instruction 1007 will result in a mis-speculation. In contrast, the invention of the '752 patent incrementally increases the prediction for each mis-speculation associated with an instruction pair, while it decrements the prediction associated with a pair of instructions when they *do not mis-speculate*, thereby updating its assessment of the likelihood of mis-speculations in the future.

Steely's tag replacement system, even as explained by Dr. Colwell, *discards* the prediction associated with a pair of instructions when a different pair of instructions mis-

F.3d at 983). What matters for purposes of anticipation is what the patent *actually* discloses, not what the inventor says it would do in a situation the patent does not clearly address.

speculates. While this data elimination admittedly yields a *change* in the result of the tag comparison, that change has nothing to do with updating the likelihood that the first pair of instructions will mis-speculate again in the future. Accordingly, no reasonable jury could find that the Steely patent discloses each and every limitation of the '752 patent as properly construed. The court will, therefore, grant summary judgment to WARF on defendant's Steely anticipation defense and counterclaim.

B. Indefiniteness

Finally, Apple contends that claims 5 and 6 of the '752 patent are invalid as indefinite. "[A] patent is invalid for indefiniteness if its claims, read in light of the specification delineating the patent, and the prosecution history, fail to inform, with reasonable certainty, those skilled in the art about the scope of the invention." *Nautilus, Inc. v. Biosig Instruments, Inc.*, 134 S. Ct. 2120, 2124 (2014). A party raising an indefiniteness challenge, like other invalidity challenges, bears the burden of proving that invalidity by clear and convincing evidence. *Microsoft Corp. v. i4i Ltd. P'ship*, 131 S. Ct. 2238, 2242 (2011); *see also* 35 U.S.C. § 282.

Here, Apple contends that claims 5 and 6 of the '752 patent should be held indefinite under *Nautilus* solely because certain terms in those claims lack an antecedent basis. Claim 5 is a dependent claim and reads:

> The data speculation decision circuit of claim 2 wherein the instruction synchronization circuit includes a synchronization table associating *the certain data consuming instructions* and *the certain data producing instructions* each with a flag value indicating whether the respective certain data producing instruction has been executed and wherein the instruction

> synchronization circuit delays the particular data consuming instruction only:
>
> i) when the prediction associated with the data consuming instruction is within a predetermined range; and
>
> ii) when the particular data consuming instruction is in *the prediction table*; and
>
> iii) when the flag indicates the particular data producing instruction has not been executed.

('752 patent, 15:7-20 (emphasis added).) Similarly, Claim 6 likewise depends from claim 2 and reads:

> The data speculation decision circuit of claim **2** wherein the instruction synchronization circuit creates an entry in *the synchronization table* including the particular data consuming instructions and data producing instructions and *the flag value* only after a mis-speculation indicating is received for the particular data consuming instruction and the particular data producing instruction.

('752 patent, 15:21-27 (emphasis added).) Apple focuses on the italicized portions of each of the above claims in making its § 112 argument.

According to Apple, the use of the definite article "the" in each of the above italicized instances suggests that the terms that article introduces must refer to specific claim elements already previously discussed. (Def.'s Br. Support Summ. J. (dkt. #118) 39.) *See also, e.g.*, *Warner-Lambert Co. v. Apotex Corp.*, 316 F.3d 1348, 1356 (Fed. Cir. 2003) ("[I]t is a rule of law well established that the definite article 'the' particularizes the subject which it precedes. It is a word of limitation as opposed to the indefinite or generalizing force of 'a' or 'an.'") (quoting *Am. Bus Ass'n v. Slater*, 231 F.3d 1, 4-5 (D.C. Cir. 2000)). As Apple points out, the italicized terms above do *not* appear elsewhere in

claims 5 and 6 themselves, or in claims 1 and 2, on which both claims 5 and 6 ultimately depend. In Apple's view, this makes it impossible for a person of skill in the art to determine the scope of claims 5 and 6, rendering them indefinite.

In *Halliburton Energy Services, Inc. v. M-I LLC*, 514 F.3d 1244 (Fed. Cir. 2008), the Federal Circuit held that "a claim could be indefinite if a term does not have proper antecedent basis where such basis is not otherwise present by implication or the meaning is not reasonably ascertainable." *Id*. at 1249; *see also Energizer Holdings, Inc. v. Int'l Trade Comm'n*, 435 F.3d 1366, 1370 (Fed. Cir. 2006) (citing *Slimfold Mfg. Co. v. Kinkead Indus., Inc.*, 810 F.2d 1113, 1116 (Fed. Cir. 1987)). The specification can, however, provide sufficient context for a person skilled in the field of the art to understand the claim to render it definite. *See, e.g.*, *In re Skvorecz*, 580 F.3d 1262, 1268 (Fed. Cir. 2009) ("We agree with Mr. Skvorecz that the clause 'welded to said wire legs at the separation' does not require further antecedent basis in claim 1, for a person skilled in the field of the invention would understand the claim when viewed in the context of the specification."). Here, the terms in question are "reasonably ascertainable" in light of the patent's specification.

Taking first the terms "the certain data consuming instructions" and "the certain data producing instructions" in claim 5, the patent's specification summarizes the invention and notes that the invention's instruction synchronization circuit:

> may also include a synchronization table associating *certain data consuming instructions and certain data producing instructions*, each with a flag indicating whether the respective data producing instruction has been executed. The instruction synchronization circuit delays the subsequent instances of the certain data consuming instruction only when the prediction

> associated with the data consuming instruction is within a predetermined range and when the particular data consuming instruction is in the prediction table and when the flag indicates that particular data producing instruction has not been executed.

('752 patent, 4:54-65 (emphasis added).) As WARF points out, this portion of the specification tracks the language of claim 5 almost exactly. There is no reason why a person of ordinary skill in the art would not read "the certain data consuming instructions" and "the certain data producing instructions" to be those included in the synchronization table in light of the specification. At the very least, the brief summary of the invention allows one skilled in the art to proceed with "reasonable certainty," as *Nautilus* requires.

The term "the prediction table" in subsection (ii) of claim 5 would likewise inform a person of ordinary skill in the art that the "prediction table" is contained in the instruction synchronization circuit. As the brief summary of the invention states, "[t]he instruction synchronization circuit may include a *prediction table* listing certain data consuming instructions and certain data producing instructions each associated with a prediction." ('752 patent, 4:39-42 (emphasis added).) The instruction synchronization circuit then employs the entries in that prediction table in determining whether to delay subsequent instances of the data consuming instruction -- the instruction must be in the prediction table for delay to take place. (*Id.* at 4:48-53.)

As for claim 6, the "synchronization table" is the one that "may" be included in the instruction synchronization circuit (which is explicitly claimed in independent claim 2) and "associate[s] certain data consuming instructions and certain data producing

instructions, each with a flag indicating whether the respective data producing instruction has been delayed." ('752 patent, 4:54-58.) The "flag value" likewise takes its meaning from this portion of the specification, which indicates that each pair of instructions in the synchronization table has "a flag indicating whether the respective data producing instruction has been executed." (*See id.*) The invention then uses the flag to determine when to delay execution of subsequent instances of the data consuming instruction. (*Id.* at 4:58-65.) A person of ordinary skill in the art would understand the scope of "the flag value" in claim 6 in light of this relatively clear context. (*Id.*)

Importantly, because Apple does not dispute that the specification offers context for the claim terms it identifies, that argument is waived. *See Fresenius USA, Inc. v. Baxter Int'l, Inc.*, 582 F.3d 1288, 1296 (Fed. Cir. 2009) ("If a party fails to raise an argument before the trial court, or presents only a skeletal or undeveloped argument to the trial court, we may deem that argument waived on appeal."); *Jordan v. Binns*, 712 F.3d 1123, 1134 (7th Cir. 2013) (undeveloped arguments considered waived); *Ultratec, Inc. v. Sorenson Commc'ns, Inc.*, No. 13-cv-346-bbc, 2014 WL 3565409, at *1 (W.D. Wis. July 17, 2014). Regardless, Apple takes an entirely different tack, one which requires a bit of explanation. According to Apple, in light of the antecedent basis problems in claims 5 and 6, a person of ordinary skill in the art *might* simply look to the specification to understand the scope of the invention. However, Apple argues, she might also assume that claims 5 and 6 do not, in fact, depend from claim 2 but instead were intended to depend from claims 3 and 5, respectively, which would provide the requisite antecedent

basis for the identified terms, but would also include additional limitations by virtue of depending from different claims. (Def.'s Br. Opp'n Summ. J. (dkt. #140) 41.)

The court does not find Apple's argument persuasive. Apple cites no cases in which courts found indefiniteness due solely to a lack of antecedent basis, at least where the specification so clearly delineates the structure of what the patent intended to claim. Instead, Apple cites *Novo Industries, L.P. v. Micro Molds Corp.*, 350 F.3d 1348 (Fed. Cir. 2003), for the proposition that claims are indefinite where "in light of the mistakes in the claims there is no clear choice as to how to interpret their scope." (Def.'s Br. Opp'n Summ. J. (dkt. #140) 46.) But *Novo* involved an obvious typographical error amenable to no fewer than four possible interpretations (at least one of which would have significant substantive implications for the scope of the claims).[16] *Novo* does not support this court *reading in* a typographical error to *create* ambiguity where the specification otherwise indisputably provides context to delineate the scope of the invention "with reasonable certainty." *Nautilus*, 134 S. Ct. at 2124.

The other case Apple cites, *Automed Technologies, Inc. v. Microfil, LLC*, 244 F. App'x 354 (Fed. Cir. 2007), is similarly unhelpful to its indefiniteness argument. In *Automed*, the Federal Circuit vacated and remanded a grant of summary judgment of non-

[16] In *Novo,* the claim included a "stop means formed on a rotatable with said support finger." 350 F.3d at 1352 (emphasis removed). Novo suggested correcting the claim either by deleting the words "a rotatable with" or by deleting the words "with said." *Id.* at 1357. The district court raised another possibility by changing the word "a" to "and." *Id.* And Micro Molds proposed as a fourth possibility that a word, such as "skirt" or "disk," might have been erroneously omitted, which would add an additional substantive limitation to the claims. *Id.* Because the Federal Circuit "[could not] know what correction [was] necessarily appropriate or how the claim should be interpreted," it concluded that the claim was necessarily indefinite "in its present form." *Id.* No comparable indefiniteness is even arguable in this case.

infringement because the district court had based its ruling on a finding that the accused systems lacked a "controller" -- a limitation that was actually *absent* from the asserted claims. *Id.* at 359. In the midst of that discussion, the Federal Circuit observed:

> We also note that claim 27 of the '671 patent, which recites "the controller," appears to be mistakenly dependent on claim 20, in which this term finds no antecedent basis. . . . Because claim 21 - and not claim 20 - recites a "controller" limitation, perhaps claim 27 was intended to depend from claim 21.

Id. Even so, the Federal Circuit said nothing about that potential error rendering claim 27 indefinite. Rather, it "[left] to AutoMed any corrective action it deem[ed] necessary." *Id.* The Federal Circuit's observation that claim 27 might have been intended to depend from claim 21, not claim 20, certainly does not compel, or even do much to support, a finding of indefiniteness in this case.

Accordingly, the court finds that the specification provides ample guidance as to what elements the claims are referencing when they refer to "the certain data consuming instructions," "the certain data producing instructions" and "the prediction table" (claim 5), as well as "the synchronization table" and "the flag value" (claim 6). Even the authority upon which defendants rely indicates that a lack of antecedent basis renders a claim indefinite only if "it would be unclear as to what element the limitation was making reference." Manual of Patent Examining Procedure § 2173.05(e) (9th ed. 2014); *see also Halliburton*, 514 F.3d at 1249. That is simply not the case here.

III. Willful Infringement

WARF has alleged a claim that Apple's infringement was willful, thereby permitting (but not requiring) the court to award enhanced damages. 35 U.S.C. § 284 ("[T]he court may increase the damages up to three times the amount found or assessed."); *Beatrice Foods Co. v. New Eng. Printing & Lithographing Co.*, 923 F.2d 1576, 1578 (Fed. Cir. 1991) ("It is well-settled that enhancement of damages must be premised on willful infringement or bad faith.") (citations omitted). Apple seeks summary judgment on this claim on the basis that WARF cannot as a matter of law meet the threshold for proving willfulness on an objective basis.

To establish willful infringement, WARF "must show by clear and convincing evidence" (1) that "the infringer acted despite an objectively high likelihood that its actions constituted infringement of a valid patent," and (2) that "this objectively-defined risk . . . was either known or so obvious that it should have been known to the accused infringer." *In re Seagate Tech.*, 497 F.3d at 1371. The former "objective determination of recklessness" is a question for the court, not the jury. *Bard Peripheral Vascular, Inc. v. W.L. Gore & Assocs., Inc.*, 682 F.3d 1003, 1006-07 (Fed. Cir. 2012).

"[T]he 'objective' prong of *Seagate* tends not to be met where an accused infringer relies on a reasonable defense to a charge of infringement." *Id.* at 1005-06 (internal citation and quotation marks omitted); *see also Spine Solutions, Inc. v. Medtronic Sofamor Danek USA, Inc.*, 620 F.3d 1305, 1319 (Fed. Cir. 2010) (overturning jury's finding of willful infringement, finding that defendant raised a "substantial question as to the obviousness" of the patent in suit); *Douglas Dynamics,* 747 F. Supp. 2d at 1112 (granting

summary judgment on willful infringement claim where there was "reasonable difference of opinion" and a "close question").

In cursory fashion, Apple's opening brief advances a wide range of arguments for seeking summary judgment on this objective prong. Some of the bases were fully briefed for review on the merits – namely, Apple's claim construction of "prediction," its related argument on anticipation by Steely and its indefiniteness defense and counterclaim as to claims 5 and 6. The court will take up Apple's motion on these bases in the discussion below.

Other bases, including ones on which Apple bears the burden of proof like obviousness, were not, however, the subject of the parties' motions for summary judgment. While the court appreciates that it is WARF's burden to demonstrate that Apple's defenses to infringement or claims of invalidity are not objectively reasonable, Apple's scattershot approach in its motion renders the task near impossible to resolve on summary judgment. Perhaps if Apple had identified two or three of its strongest arguments, this may have been a manageable task. Instead, Apple's treatment of each basis is limited to a paragraph or two in its opening brief and reflects ships passing in the night in reply to WARF's responses.[17] In any event, WARF did come forward with evidence and law that, despite Apple's attempt to refute it in reply, could lead to a finding that Apple's belief that it either did not infringe the '752 patent or that the

[17] Perhaps most telling, the few defenses that Apple moved on the merits do not offer grounds for the court to find for Apple on the objective prong of WARF's willful infringement claim.

patent was invalid was not objectively reasonable.[18] As such, the court will await a more robust demonstration of the merits of Apple's defenses and WARF's infringement claims at trial.[19]

Returning to those bases which *were* fully briefed for review on the merits, Apple's claim construction is arguably "objectively reasonable" if viewed purely in a vacuum. Apple presented some evidence that "prediction" *can* describe a static value in the context of computer circuits and speculation, for example, in the form of the Smith article and Chrysos paper; they are also correct that the patent does not *explicitly* define "prediction," ostensibly leaving at least some room for debate. The problem is that *nothing* in the patent -- not the claim language, not the specification, not the purpose of the invention -- supports Apple's construction. As discussed above, the claim language from the outset suggests that a prediction must be dynamic in the context of this particular invention. The specification, including both the brief summary of the invention and the detailed description of the embodiments, further supports this construction. And Apple's resort to extrinsic evidence fails to render its arguments to the contrary any more reasonable,

---

[18] Certainly, Apple seems to make an objectively reasonable argument as to claims 1 and 2 being obvious, and perhaps as to claims 3 and 9, but the court cannot say on this record whether the supposed links drawn between Steely, Hesson, Chen and EV6 are obvious or pure sophistry. Similarly, while Apple raises a number of arguments that appear to objectively establish non-infringement on a literal basis, it has left the court unconvinced as to infringement under the doctrine of equivalents.

[19] To clarify, while the jury is deliberating on liability, the court can take up any additional evidence and argument relevant to the objective prong of WARF's willful infringement claim and likely will render a decision on the objective prong before the parties present any evidence on the subjective prong during the second phase of the trial (assuming the jury finds infringement and does not find invalidity).

given that extrinsic evidence cannot be used to vary the intrinsic evidence under settled principles of claim construction.

While superficially appealing, not unlike a siren's song, Apple's construction crashes against the rocks of the patent language itself and intrinsic evidence. Given how strongly the patent itself supports WARF's narrower construction, and how little Apple has to offer in support of its broader one, Apple's position is not objectively reasonable. *Compare Cohesive Techs., Inc. v. Waters Corp.*, 543 F.3d 1351, 1374 (Fed. Cir. 2008) (no willfulness where disputed term "was susceptible to a reasonable construction under which [the] products did not infringe"), *with SSL Services, LLC v. Citrix Systems, Inc.*, 769 F.3d 1073, 1091 (Fed. Cir. 2014) (affirming district court's finding of willful infringement, in part, because defendant's non-infringement defense based on an unwarranted limitation of a claim term was not objectively reasonable); *cf. Raylon, LLC v. Complus Data Innovations, Inc.*, 700 F.3d 1361, 1369 (Fed. Cir. 2012) (finding position on claim construction frivolous under Rule 11 where proffered construction was "contrary to all the intrinsic evidence and does not conform to the standard canons of claim construction"). Finally, while Judge Crabb granted summary judgment to the defendant in *Intel* on WARF's willful infringement claim, she did so on a basis unrelated to claim construction and one not before this court. *Intel*, 656 F. Supp. 2d at 924 (finding defendant's licensing defense objectively reasonable). Accordingly, the court will deny defendant's motion for summary judgment on plaintiff's willful infringement claim that depends on Apple's claim construction, finding this defense objectively unreasonable.

As for Apple's anticipation challenge to the validity of the '752 patent based solely on Steely, the court finds this defense not objectively reasonable as well, though it will reserve on any obviousness defense involving Steely. Much of Apple's anticipation argument depended upon its claim construction, which was not objectively reasonable as discussed above. Admittedly, Apple attempted to maintain an anticipation defense even under WARF's construction, but its dependence on Steely's purported, defective "tag overwriting" scheme is likewise unreasonable, given that this dubious overwriting defect would simply *dispose* of previous predictions, rather than "updating" them to reflect an increased likelihood of future mis-speculation.

The court will also deny Apple's motion with respect to its indefiniteness defense. Apple points to *no* case in which a lack of antecedent basis led to a finding of indefiniteness *despite* clear context providing that basis in the specification. Even the case law Apple cites explain that there is no invalidity for indefiniteness so long as the antecedent basis is present by implication, and Apple waived any contention that the specification did *not* serve to provide such context. As a whole then, this defense was not objectively reasonable, and Apple cannot use it to escape the possibility of enhanced damages.

ORDER

IT IS ORDERED that:

1) Plaintiff Wisconsin Alumni Research Foundation's construction of the disputed term "prediction" is ADOPTED as described in this opinion.

2) Defendant and counter claimant Apple, Inc.'s motion for summary judgment (dkt. #116) is DENIED as to its counterclaims and defenses of anticipation by Steely and indefiniteness, and DENIED as to plaintiff's willful infringement claim premised on (1) Apple's claim construction, (2) anticipation by Steely, and (3) indefiniteness of claims 5 and 6. The court RESERVES on the motion in all other respects.

3) Plaintiff's motion for summary judgment (dkt. #117) is GRANTED.

Entered this 5th day of August, 2015.

BY THE COURT:

/s/

WILLIAM M. CONLEY
District Judge